

For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex LIBRIS
UNIVERSITATIS
ALBERTAENSIS





Digitized by the Internet Archive
in 2024 with funding from
University of Alberta Library

<https://archive.org/details/Larsson1973>

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Stig. O. LARSSON

TITLE OF THESIS EDF - Statistics in the case of
..... exponential distribution
.....

DEGREE FOR WHICH THESIS WAS PRESENTED M.Sc. - Mathematical Statistics

YEAR THIS DEGREE GRANTED Spring - 1973

Permission is hereby granted to THE UNIVERSITY OF ALBERTA
LIBRARY to reproduce single copies of this thesis and to lend
or sell such copies for private, scholarly or scientific research
purposes only.

The author reserves other publication rights, and neither
the thesis nor extensive extracts from it may be printed or
otherwise reproduced without the author's written permission.

THE UNIVERSITY OF ALBERTA

EDF-STATISTICS IN THE CASE OF EXPONENTIAL DISTRIBUTION

by



STIG O. LARSSON

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

IN STATISTICS

DEPARTMENT OF MATHEMATICS

EDMONTON, ALBERTA

SPRING, 1973

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled EDF-STATISTICS IN THE CASE OF EXPONENTIAL DISTRIBUTION submitted by Stig O. LARSSON in partial fulfilment of the requirements for the degree of Master of Science in Statistics.

ABSTRACT

In this thesis we considered some tests which may be used to decide whether or not a given sample comes from exponential distribution with an unknown parameter.

We first justified usage of the standard EDF-statistics to test for exponentiality and secondly we did a power study for different tests, by the Monte Carlo method. This is a part of a study done by M. A. Stephens, for usage of the EDF-statistic for goodness-of-fit in general.

ACKNOWLEDGEMENTS

I wish to thank Dr. S. G. Ghurye for his constant supervision and Dr. M. Stephens for suggesting the problem and for his help with the work. I also wish to thank Dr. E. S. Keeping for proof-reading my thesis.

TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGEMENTS	(i)
TABLE OF CONTENTS	(ii)
CHAPTER I	
Introduction	1
EDF-Statistics	3
Some Practical Simplifications	6
Studentized Integral Transformation	8
CHAPTER II	
Introduction	20
Properties of the Exponential Distribution	21
Transformation from Exponential to Uniform Distribution	32
The W-Statistics	40
CHAPTER III	
Introduction	47
Modification of EDF-Statistics	47
Chi-Square Approximation	49
TABLE 1	52
TABLE 2	53
TABLE 3	54
TABLE 4A	55
TABLE 4B	56
TABLE 4C	57
REFERENCES	58

CHAPTER I

1.1.0 Introduction.

An important problem in statistics relates to obtaining information about the form of the population from which the sample is drawn. This problem is known as the goodness-of-fit problem. Stated in more exact terms we have the following: Let $\{X_i \mid 1 \leq i \leq n\}$ be independent, identically distributed (i.i.d.) random variables (r.v.'s) with the distribution function F . The goodness-of-fit problem is to devise a test of the hypothesis

$$H_0 : F(x) = G(x)$$

where $G(x)$ is a given distribution function.

The problem can be divided into two parts:

- (a) $G(x)$ is completely specified, for example, it may be the exponential distribution with mean 1 on the non-negative reals, or
- (b) $G(x)$ is known to be a member of a specified family, for example, it may be an exponential distribution on the non-negative reals with an unspecified mean.

The classical test for the goodness-of-fit problem is the chi-square statistic, which has several advantages over the more modern EDF-statistics (so called because the statistics are

based on the empirical distribution function). For example, the chi-square statistic can be used for both continuous and discrete cases. It is also known how to use the chi-square statistic in both cases (a) and (b). However, in general the EDF-statistics give more powerful tests of H_0 than the chi-square and in at least the case when the hypothesised distribution is Normal or Exponential, the EDF-statistic can be used in case (b). In section 1.2.0 we will define the most common of the EDF-statistics, namely, the Cramér-von Mises, the Anderson-Darling, the Watson, the Kolmogorov-Smirnov and the Kuiper statistics.

In section 1.3.0 we will show that the EDF-statistics can be expressed by some simple formulas for purposes of calculation. In so doing it becomes clear that the tests using these statistics are equivalent to testing whether $\{X_i \mid 1 \leq i \leq n\}$ has a distribution G (completely specified) or to testing if $\{F(X_i) \mid 1 \leq i \leq n\}$ has a uniform $(0,1)$ distribution.

If G is not completely specified, we must estimate the parameters $\theta_1, \dots, \theta_s$ (say) by $T_1 = t_1(X_1, \dots, X_n)$, $\dots$, $T_s = t_s(X_1, \dots, X_n)$ (denoted collectively by $\underline{T}$) and then make the transformation

$$Z_i = G(X_i \mid \underline{T}) \quad i = 1, \dots, n \quad .$$

We find the Z_i 's to be neither independent nor uniformly distributed, and in general the distribution of Z_i will depend

on G as well as on the true parameters $\underline{\theta}$. This problem we will discuss in section 1.4.0.

1.2.0 EDF-Statistics.

From now on we will only consider r.v.'s with a density function.

Let $\{X_i \mid 1 \leq i \leq n\}$ be i.i.d. r.v.'s with distribution function F . Let

$$\varepsilon(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

then the empirical distribution function based on the sample $\{x_i \mid 1 \leq i \leq n\}$ is defined by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \varepsilon(x - X_i) ,$$

and hence $F_n(x)$ is the proportion of the X_i , $i = 1, \dots, n$, which are less than or equal to x .

The EDF-statistics are based on the differences between $F_n(x)$ and the hypothesised distribution F . For this to make sense we need $F_n(x)$ to converge to F , in some sense, under the null-hypothesis. Hence the following theorem:

1.2.1 Theorem. (Glivenko-Cantelli Theorem).

As $n \rightarrow \infty$

$$\sup_{-\infty < x < \infty} |F_n(x) - F(x)| \rightarrow 0 \quad .$$

(We omit the proof. For reference see, for example, K.L. Chung "A course in probability theory".)

In 1928 Cramér suggested the following test function for H_0

$$\int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 dK(x)$$

where $K(x)$ is a suitable non-decreasing weight function. Von-Mises independently made an equivalent suggestion. Smirnov [16] gave the modification

$$C_n^2 = n \int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 \Psi(F(x)) dF(x)$$

where $\Psi(t)$, $0 \leq t \leq 1$ is a non-negative weight function. In case $\Psi(t) \equiv 1$, we will call this the Cramér-con Mises statistic and denote it by W_n^2 i.e.

$$W_n^2 = n \int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 dF(x) \quad .$$

The exact distribution of W_n^2 is unknown; however, the limit distribution as $n \rightarrow \infty$ was found by Smirnov.

Anderson and Darling [1] considered the weight function $\Psi(t) = \frac{1}{t(1-t)}$. In this case we will use the notation

$$A_n = n \int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 \frac{1}{F(x)(1-F(x))} dF(x)$$

and call it the Anderson-Darling statistic. Anderson and Darling also found the limiting distribution of A_n , and Marshall [10] showed that the asymptotic critical values are valid for $n \geq 5$ for all practical purposes.

It is clear that the Anderson-Darling statistic is more sensitive to differences in the "tails" between $F_n(x)$ and $F(x)$ than the Cramér-von Mises statistic since $\psi(t) = \frac{1}{t(1-t)}$ is minimum for $t = \frac{1}{2}$ and increases as $t \rightarrow 0$ or 1 .

Another variation of the Cramér statistic was introduced by Watson [18], and is defined by

$$U_n^2 = n \int_{-\infty}^{+\infty} (F_n(x) - F(x)) - \int_{-\infty}^{+\infty} (F_n(x) - F(x)) dF(x))^2 dF(x) .$$

We will call this the Watson statistic. Watson also found the limiting distribution for U_n^2 .

In 1933 Kolmogorov [6] suggested a test of H_0 based on the statistic

$$K_n = \sqrt{n} \sup_{-\infty < x < \infty} |F_n(x) - F(x)| .$$

He also found the limiting distribution of K_n . Here we will use the modifications

$$D_n^+ = \sup_{-\infty < x < \infty} \{F_n(x) - F(x)\}$$

$$D_n^- = \sup_{-\infty < x < \infty} \{F(x) - F_n(x)\}$$

$$D_n = \max (D_n^+, D_n^-) \quad .$$

They are called the Kolmogorov-Smirnov statistics. Smirnov [15] simplified the derivation of the limit distribution, and also tabulated D_n .

Kuiper [7] suggested the modification

$$V_n = D_n^+ + D_n^- \quad .$$

We will call this the Kuiper statistic. The distribution function for $\sqrt{n} V_n$ was also found by Kuiper.

1.3.0 Some Practical Simplifications.

The definitions in section 1.2.0 can be expressed for computation more simply as in lemma 1.3.1. The formulas in 1.3.1 for W_n^2 , A_n and U_n^2 can be obtained by straight-forward but tedious integration and the rest follow directly since $F(x)$ is monotone increasing. Therefore, the proof is omitted.

1.3.1 Lemma.

Let $\{X_i \mid 1 \leq i \leq n\}$ be i.i.d r.v.'s with the common continuous distribution function $F(x)$. Let $\{X_{(i)} \mid 1 \leq i \leq n\}$ denote the order statistics of $\{X_i \mid 1 \leq i \leq n\}$ with $X_{(1)} < X_{(2)} < X_{(n)}$. Then

$$(i) \quad W_n^2 = \sum_{i=1}^n (F(X_{(i)}) - \frac{2i-1}{2n})^2 + \frac{1}{12n}$$

$$(ii) \quad A_n = -\frac{1}{n} \left[\sum_{i=1}^n (2i-1) \{ \ln(F(X_{(i)})) + \ln(1-F(X_{n+1-i})) \} \right] - n$$

$$(iii) \quad U_n^2 = W_n^{2-n} \left(\frac{\sum_{i=1}^n F(X_{(i)})}{n} - \frac{1}{2} \right)^2$$

$$(iv) \quad D_n^+ = \max_{1 \leq i \leq n} \left(\frac{i}{n} - F(X_{(i)}) \right)$$

$$(v) \quad D_n^- = \max_{1 \leq i \leq n} \left(F(X_{(i)}) - \frac{i-1}{n} \right)$$

$$(vi) \quad D_n = \max(D_n^+, D_n^-)$$

$$(vii) \quad V_n = D_n^+ + D_n^-$$

where W_n^2 , A_n , etc., are all defined in section 1.2.0.

For each of these statistics the region of rejection for testing H_0 consists of all sufficiently large values of the statistic.

1.3.2 Lemma.

Let $\{X_i \mid 1 \leq i \leq n\}$ be i.i.d. r.v.'s with density function f , then $\{Z_i \mid 1 \leq i \leq n\}$ defined by

$$Z_i = \int_{-\infty}^{X_i} f(t) dt = F(X_i)$$

are independent with uniform $(0,1)$ distribution.

Hence $F(X_{(i)})$, $i = 1, \dots, n$, behaves as an ordered sample from uniform $(0,1)$ distribution under the null-hypothesis, and therefore the distribution EDF-statistics in lemma 1.3.1 are all independent of the distribution of the X_i 's .

1.4.0 Studentized Integral Transformation.

Let $\{X_i \mid 1 \leq i \leq n\}$ be i.i.d. r.v.'s with density function $f(x, \underline{\theta})$, θ_i unknown $i = 1, \dots, s$. Let θ_i be estimated by $T_i = t_i(\underline{X})$ for $i = 1, \dots, s$. Then the studentized integral transformation we define by

$$Z = \int_{-\infty}^X f(u, \underline{T}) du = F(X, \underline{T}) \quad .$$

In general the distribution function of Z will depend on F , and the true parameters $\underline{\theta}$.

If it so happens that the joint distribution function of $Z_i = F(X_i, \underline{T})$, $i = 1, \dots, n$ does not depend on the parameters $\underline{\theta}$, we define the studentized Cramér-von Mises statistic by

$$\tilde{W}_n^2 = \sum_{i=1}^n (F(X_{(i)}, \underline{T}) - \frac{2i-1}{2n})^2 + \frac{1}{12n} \quad .$$

It is clear that the distribution of $\tilde{W}_n^2$ does not depend on $\underline{\theta}$, but will depend on the family of distributions of X_i .

Hence the following relation exists between W_n^2 and $\tilde{W}_n^2$:

$$\begin{aligned} \tilde{W}_n^2 = W_n^2 + 2 \sum_{j=1}^n \{ (F(X_{(j)}, \underline{\theta}) - \frac{2j-1}{2n}) (F(X_{(j)}, \underline{T}) - F(X_{(j)}, \underline{\theta})) \} \\ + \sum_{j=1}^n (F(X_{(j)}, \underline{T}) - F(X_{(j)}, \underline{\theta}))^2 . \end{aligned}$$

Similarly we can define studentized EDF-statistics for the Anderson-Darling statistic, the Watson statistic, etc. We will denote the studentized EDF-statistics by $\tilde{A}_n$, $\tilde{U}_n^2$, etc.

1.4.1 The Joint Density of $\{Z_i\}$.

Let

$$T_r = t_r(\underline{X}) \quad , \quad r = 1, \dots, s$$

$$Z_i = F(X_i, \underline{T}) \quad , \quad i = 1, \dots, n .$$

We assume T_r and Z_i are differentiable with respect to X_j , for all r, i and j .

Let A denote the $(n \times n)$ matrix defined by

$$A = (a_{ij})$$

where

$$a_{ij} = \frac{\partial Z_i}{\partial X_j} \quad , \quad i, j = 1, \dots, n .$$

If the transformation from x_i to z_i is one-to-one and

$|A| \neq 0$ for $\{X \mid \prod_{i=1}^n f(x_i) > 0\}$ then the joint density of $z_1, z_2, \dots, z_n$ is given by

$$P(z_1, \dots, z_n) = \prod_{i=1}^n f(x_i) \cdot |A|^{-1}.$$

If $|A| \neq 0$ this implies that the rank of $A(\rho(A))$ is equal to n . This is clearly not always so since we must estimate the s parameters in the transformation, it is neither true that the rank of A is equal to $n-s$ in general.

So the problem is first to find the rank of A . This can be done as follows:

Let

$$g_{io} = f(x_i, \underline{T})$$

$$g_{ir} = \frac{\partial}{\partial t_r} F(x_i, \underline{T})$$

$$t_{ri} = \frac{\partial t_r}{\partial x_i}$$

$$i = 1, \dots, n, \quad , \quad r = 1, \dots, n.$$

Then

$$\frac{\partial z_i}{\partial x_j} = \delta_{ij} g_{io} + \sum_{r=1}^s g_{ir} t_{rj}$$

where

$$\delta_{ij} = \begin{cases} 0 & , \quad i \neq j \\ 1 & , \quad i = j \end{cases}.$$

Let T be the $(s \times n)$ matrix, defined by

$$T = (t_{ri})$$

$$r = 1, \dots, s, \quad i = 1, \dots, n$$

and G the $(n \times s)$ matrix defined by

$$G = \left(\frac{g_{ir}}{g_{i0}} \right), \quad g_{i0} > 0$$

$$r = 1, \dots, s, \quad i = 1, \dots, n.$$

Then we may write $|A|$ as

$$|A| = \prod_{i=1}^n g_{i0} |I_n + GT|$$

where I_n is the $(n \times n)$ identity matrix. Clearly

$$\rho(A) = \rho(I_n + GT).$$

Consider the partitioned $((n+s) \times (n+s))$ matrix

$$B = \left(\begin{array}{c|c} I_n + GT & 0 \\ \hline 0 & I_s \end{array} \right).$$

We have $\rho(A) = \rho(B) - s$. But

$$\left(\begin{array}{c|c} I_n & 0 \\ \hline T & I_s \end{array} \right) \left(\begin{array}{c|c} I_n & G \\ \hline 0 & I_s \end{array} \right) \left(\begin{array}{c|c} I_n + GT & 0 \\ \hline 0 & I_s \end{array} \right) \left(\begin{array}{c|c} I_n & 0 \\ \hline -T & I_s \end{array} \right) \left(\begin{array}{c|c} I_n & -G \\ \hline 0 & I_s \end{array} \right) = \left(\begin{array}{c|c} I_n & 0 \\ \hline 0 & I_s + TG \end{array} \right) = C.$$

Clearly $\rho(B) = \rho(C) = \rho(I_S + TG) + n$.

Hence we have proved the following lemma:

1.4.2 Lemma.

$$\rho(A) = n - s + \rho(I_S + TG).$$

1.4.3 Example.

Let $\{X_i \mid 1 \leq i \leq n\}$ be i.i.d. r.v.'s with the distribution function

$$F(x) = 1 - e^{-x/\beta}$$

$x \geq 0$ and $\beta > 0$ (β unknown).

Let $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ be estimated by $\bar{X}$ and consider the transformation

$$Z_i = 1 - e^{-\frac{X_i}{\bar{X}}}, \quad i = 1, \dots, n$$

then the rank of A is equal to $(n-1)$ since

$$g_{10} = \frac{1}{\bar{x}} \exp\left\{-\frac{X_i}{\bar{X}}\right\}$$

$$g_{11} = -\frac{X_i}{\bar{x}^2} \exp\left\{-\frac{X_i}{\bar{X}}\right\}$$

$$I_S + TG = 1 + \sum_{i=1}^n \left(\frac{1}{n}\right) \left(-\frac{X_i}{\bar{x}}\right) = 0$$

and by lemma 1.4.2 the result follows.

For some of the standard distributions such as Normal or chi-square with two degrees of freedom (d.f.) the rank of A is equal to $n-s$. In this case the jacobian of the transformation can be simplified as follows:

Assume that there are $n-s$ linearly independent relationships between the variables $z_1, \dots, z_n$, say $z_1, \dots, z_{n-s}$. We also assume that the transformation

$$T_r = t_r(x) \quad , \quad r = 1, \dots, s$$

$$z_i = F(x_i, \underline{T}) \quad , \quad i = 1, \dots, n-s$$

is one-to-one. Since the determinant of every matrix is equal to the determinant of the transpose of that matrix, we have

$$|J| = \begin{vmatrix} g_{10} + \sum_{r=1}^s g_{1r} t_{r1} , \dots , \sum_{r=1}^s g_{n-sr} t_{r1} , & t_{11}, \dots, t_{s1} \\ & , g_{n-s0} + \sum_{r=1}^s g_{n-sl} t_{rn-s} , \\ \sum_{r=1}^s g_{ir} t_{rn} , & , \sum_{r=1}^s g_{n-sr} t_{rn} , & t_{1n}, \dots, t_{sn} \end{vmatrix}$$

Multiplying the matrix J by B where

$$B = \left[\begin{array}{c|c} \frac{I_{n-s}}{-g_{11}, \dots, -g_{n-s1}} & \frac{0}{I_s} \\ \hline -g_{1s}, \dots, -g_{n-s,s} & \end{array} \right]$$

it is easy to see that

$$|J| = \prod_{i=1}^{n-s} g_{10} \left| \frac{\partial(t_1, \dots, t_s)}{\partial(x_{n-s+1}, \dots, x_n)} \right| .$$

Hence the joint density of $z_1, \dots, z_{n-s}, T_1, \dots, T_s$ can be written as

$$P(z_1, \dots, z_{n-s}, T_1, \dots, T_s) = \frac{\prod_{i=1}^n f(x_i, \underline{\theta})}{\prod_{i=1}^{n-s} f(x_i, \underline{T})} \left| \frac{\partial(t_1, \dots, t_s)}{\partial(x_{n-s+1}, \dots, x_n)} \right|^{-1} .$$

An alternative expression for the joint density may be obtained by letting

$$\prod_{i=1}^{n-s} f(x_i, \underline{\theta}) = P(x_1, \dots, x_{n-s} | \underline{\theta})$$

then

$$\begin{aligned} P(z_1, \dots, z_{n-s}, T_1, \dots, T_s) &= \prod_{i=1}^{n-s} \frac{f(x_i, \underline{\theta})}{f(x_i, \underline{T})} \prod_{i=n-s+1}^n f(x_i, \underline{\theta}) \\ \left| \frac{\partial(t_1, \dots, t_s)}{\partial(x_{n-s+1}, \dots, x_n)} \right| &= \prod_{i=1}^{n-s} \frac{f(x_i, \underline{\theta})}{f(x_i, \underline{T})} P(T_1, \dots, T_s, x_1, \dots, x_{n-s} | \underline{\theta}) \\ &= \frac{P(x_1, \dots, x_{n-s}, T_1, \dots, T_s | \underline{\theta})}{\prod_{i=1}^{n-s} f(x_i, \underline{T})} . \end{aligned}$$

1.4.4 Example.

Let $\{X_i \mid 1 \leq i \leq n\}$ be i.i.d. r.v.'s with distribution function $F(x) = 1 - e^{-x/\beta}$, $x \geq 0$ and $\beta > 0$. Then the joint

density function of $z_i = 1 - e^{-x_i/\bar{x}}$, $i = 1, \dots, n-1$ is given by

$$P(z_1, \dots, z_{n-1} | \beta) = \frac{(n-1)!}{n^{(n-1)} \prod_{i=1}^{n-1} (1-z_i)} e^{-n \sum_{i=1}^{n-1} (1-z_i)} < 1$$

0 otherwise .

To show this, we first find

$$P(z_1, \dots, z_{n-1}, \bar{x} | \beta) = \frac{P(x_1, \dots, x_{n-1}, \bar{x}, \beta)}{\prod_{i=1}^{n-1} f(x_i, \bar{x})}$$

and then integrate out $\bar{x}$. Now, the joint density function of $x_1, \dots, x_n$ is given by

$$P(x_1, \dots, x_n | \beta) = \beta^{-n} \exp \left\{ -\frac{1}{\beta} \sum_{i=1}^n x_i \right\} .$$

Since the jacobian of the transformation of $(x_1, \dots, x_n)$ into $(x_1, \dots, x_{n-1}, \bar{x})$ is equal to n , the joint density of $(x_1, \dots, x_{n-1}, \bar{x})$ is given by

$$P(x_1, \dots, x_{n-1}, \bar{x} | \beta) = n\beta^{-n} \exp \left\{ -\frac{n\bar{x}}{\beta} \right\}$$

and we also have

$$\prod_{i=1}^{n-1} f(x_i, \bar{x}) = \bar{x}^{-(n-1)} \exp \left\{ -\frac{\sum_{i=1}^{n-1} x_i}{\bar{x}} \right\} .$$

Hence the joint density function of $z_1, \dots, z_{n-1}, \bar{x}$ is therefore

$$\begin{aligned} P(z_1, \dots, z_{n-1}, \bar{x} | \beta) &= \frac{n\beta^{-n} \exp \left\{ -\frac{n\bar{x}}{\beta} \right\}}{\bar{x}^{-(n-1)} \exp \left\{ -\frac{\sum_{i=1}^{n-1} x_i}{\bar{x}} \right\}} \\ &= \left(\frac{n}{\beta} \right) \left(\frac{\bar{x}}{\beta} \right)^{n-1} \exp \left[-\left\{ \frac{n\bar{x}}{\beta} - \frac{\sum_{i=1}^{n-1} x_i}{\bar{x}} \right\} \right] \end{aligned}$$

since

$$z_i = 1 - e^{-\frac{x_i}{\bar{x}}}$$

we have

$$\frac{x_i}{\bar{x}} = -\ln(1 - z_i) \quad .$$

Therefore, the joint probability density function can be written as

$$P(z_1, \dots, z_{n-1}, \bar{x} | \beta) = \left(\frac{n}{\beta} \right) \left(\frac{\bar{x}}{\beta} \right)^{n-1} e^{-\frac{n\bar{x}}{\beta}} \left(\prod_{i=1}^{n-1} \frac{1}{(1 - z_i)} \right) .$$

Integrating with respect to $\bar{x}$ gives us

$$\begin{aligned}
 P(z_1, \dots, z_{n-1} | \beta) &= \int_0^\infty \left(\frac{n}{\beta}\right) \left(\frac{\bar{x}}{\beta}\right)^{n-1} e^{-\frac{n\bar{x}}{\beta}} \left(\frac{1}{\prod_{i=1}^{n-1} \pi (1-z_i)} \right) d\bar{x} = \\
 &= \left(\frac{n}{\beta}\right) \frac{1}{n^{n-1} \prod_{i=1}^{n-1} \pi (1-z_i)} \int_0^\infty \left(\frac{n\bar{x}}{\beta}\right)^{n-1} e^{-\frac{n\bar{x}}{\beta}} d\bar{x}
 \end{aligned}$$

after substituting $y = \frac{\bar{x}n}{\beta}$, we have

$$\begin{aligned}
 P(z_1, \dots, z_{n-1} | \beta) &= \frac{(n-1)!}{n^{n-1} \prod_{i=1}^{n-1} \pi (1-z_i)} e^{-n < \prod_{i=1}^{n-1} (1-z_i) < 1} \\
 &= 0 \quad \text{otherwise}
 \end{aligned}$$

Hence the joint density function of $z_1, \dots, z_{n-1}$ is independent of the parameter β .

The marginal density of z_i is easier to find directly instead of trying to integrate over $z_j, j=1, \dots, n-1, i \neq j$. To do so we first find the distribution of $U_i = \frac{x_i}{\bar{x}}$, then the result will follow trivially.

$$P(U_i \leq a) = P\left(\frac{x_i}{\bar{x}} \leq a\right) = P\left(x_i \left(1 - \frac{a}{n}\right) \leq \frac{a}{n} \sum_{\substack{j=1 \\ j \neq i}}^n x_j\right)$$

Now the characteristic function of x_j is equal to

$$(1-it\beta)^{-1}$$

Hence the characteristic function of $Y = \frac{a}{n} \sum_{\substack{j=1 \\ j \neq i}}^n x_j$ is equal to

$$(1 - \frac{i\beta a}{n}t)^{-(n-1)}$$

Similarly the characteristic function of $V = x_i(1 - \frac{a}{n})$ is equal to

$$(1 - i\beta(1 - \frac{a}{n})t)^{-1}$$

Hence $\frac{2Y_n}{\beta a}$ and $\frac{2V_n}{\beta(n-a)}$ have a gamma distribution

$$P(V \leq Y) = \int_0^\infty \int_0^y \frac{e^{-\frac{x}{\beta(1-\frac{a}{n})}}}{\beta(1-\frac{a}{n})} \cdot \frac{y^{n-2} e^{-\frac{yn}{\beta a}}}{\Gamma(n-1) (\frac{\beta a}{n})^{n-1}} dx dy.$$

$$= 1 - \int_0^\infty \frac{(\frac{yn^2}{\beta a(n-a)})^{n-2} e^{-\frac{y}{\beta}(\frac{n^2}{a(n-a)})}}{\Gamma(n-1) (\frac{\beta a}{n})^{n-1} (\frac{n^2}{\beta a(n-a)})^{n-2}} dy$$

Substitute $u = \frac{yn^2}{\beta a(n-a)}$

We have $p(U_i \leq a) = 1 - \frac{1}{(\frac{n}{n-a})^{n-1}} = 1 - (1 - \frac{a}{n})^{(n-1)}$

Hence the density of U_i is given by

$$p(U_i) = \frac{n-1}{n} (1 - \frac{U_i}{n})^{n-2} \quad 0 \leq U_i \leq n$$

$$= 0 \quad \text{otherwise}$$

Now since $z_i = 1 - e^{-U_i}$

$$U_i = -\ln(1 - z_i)$$

Hence

$$P(z_i) = \frac{n-1}{n} \frac{1}{(1-z_i)} [1 + \frac{1}{n} \ln(1-z_i)]^{n-2} \quad 0 \leq z_i \leq 1 - e^{-n}$$

$$= 0 \quad \text{otherwise}$$

Clearly $z_1, \dots, z_{n-1}$ are not independent. The density of z_i does not depend on β . This was clear since the joint density of $z_1, \dots, z_{n-1}$ did not depend on β . We also note that the density function of z_i tends to uniform $(0,1)$ as $n \rightarrow \infty$.

CHAPTER II

2.1.0 Introduction.

In this chapter we will consider the goodness-of-fit problem, when the hypothesised distribution is exponential with the scale parameter unknown, and in one case when the location parameter is also unknown.

There is a fairly extensive literature on testing whether a random sample is consistent with an exponential population. Some of the contributors to this problem are Epstein and Sobel [4], Seshadri, Csörgö and Stephens [11], Lilliefors [8] and Stephens [13].

In section 2.2.0 we will derive some of the properties of the exponential distribution such as the mean and variance of the order statistics and sufficient and complete statistics for the parameters. The properties we derive are not in any way an extensive survey of the exponential distribution, but only what is needed for deriving some methods for testing for exponentiality.

It is clear from the examples in section 1.4.0 that one way to test for exponentiality is to use the studentized integral transformation, and then use the modified EDF-statistic. In section 2.3.0 we will discuss two other methods. These are based on a transformation from exponentiality with an unknown scale parameter into a completely

specified distribution, in this case uniform. Instead of testing the original data for exponentiality we will transform the data and test for uniformity. A test to do so has already been discussed in section 1.3.0.

Shapiro and Wilks [12] introduced a statistic to test for exponentiality based on generalized least squares. This test can be used when both the location and the scale parameters are unknown. We will discuss this test in section 2.4.0.

2.2.0 Properties of the Exponential Distribution

2.2.1 Definition

A random variable X is said to have an exponential distribution if and only if (iff) its density is

$$\begin{aligned} f(x) &= \frac{1}{\sigma} \exp\left\{-\frac{x-\theta}{\sigma}\right\} & x \geq \theta \\ &= 0 & \text{otherwise} \end{aligned}$$

where the location parameter $-\infty < \theta < +\infty$ and the scale parameter $\sigma > 0$.

The exponential distribution will be denoted by $\exp(\theta, \sigma)$.

We state without proof the well-known facts:

the mean of $\exp(\theta, \sigma)$ is equal to $\theta + \sigma$, and the variance is equal to σ^2 .

The properties we are most interested concern the order statistics. Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with density function $f(x)$, then the joint density of the order statistics $(X_{(i)})_{i=1, \dots, n}$ is given by

$$g(x_{(1)}, \dots, x_{(n)}) = n! \prod_{i=1}^n f(x_{(i)}) \quad x_{(1)} < x_{(2)} < \dots < x_{(n)}$$

$$= 0 \quad \text{otherwise}$$

so clearly if $\{X_j\}$ are $\exp(\theta, \sigma)$ we have

$$g(x_{(1)}, \dots, x_{(n)}) = \frac{n!}{\sigma^n} \exp\left\{-\frac{1}{\sigma} \left(\sum_{i=1}^n (x_i - \theta)\right)\right\} \quad x_{(1)} < x_{(2)} < \dots < x_{(n)}$$

$$= 0 \quad \text{otherwise}$$

2.2.2 Lemma

Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s. Define $Y_1 = X_{(1)} - \theta$

$$Y_i = X_{(i)} - X_{(i-1)}, \quad i=2, \dots, n$$

then the Y_i 's are independent and the distribution of Y_i is given by $\exp(0, \frac{\sigma}{n+1-i})$ iff the distribution of X_i is given by $\exp(\theta, \sigma)$.

Proof

The order statistics of $\{X_i\}$ have the density

$$g(x_{(1)}, \dots, x_{(n)}) = n! \sigma^{-n} \exp\left\{-\frac{1}{\sigma} \left(\sum_{i=1}^n (x_i - \theta) \right)\right\} \quad x_{(1)} < x_{(2)} < \dots < x_{(n)}$$

$$= 0 \quad \text{otherwise}$$

The jacobian of the transformation of X_i to Y_i is equal to 1, therefore, the density of $Y_1, \dots, Y_n$ is given by

$$p(y_1, \dots, y_n) = n! \sigma^{-n} \exp\left\{-\frac{1}{\sigma} (y_1 + y_2 + y_1 + \dots + y_n + y_{n-1} \dots + y_1)\right\}$$

$$= n! \sigma^{-n} \exp\left\{-\frac{1}{\sigma} [ny_1 + (n-1)y_2 + \dots + y_n]\right\}$$

Since the joint density can be written as the product of $y_1, \dots, y_n$ i.e.

$$p(y_1, \dots, y_n) = \frac{n}{\sigma} \exp\left\{-\frac{ny_1}{\sigma}\right\} \times \frac{(n-1)}{\sigma} \exp\left\{-\frac{(n-1)y_2}{\sigma}\right\} \times \dots \times \frac{1}{\sigma} \exp\left\{-\frac{y_n}{\sigma}\right\}.$$

$y_1, \dots, y_n$ are independent, and the distribution of y_1 is $\exp(0, \frac{\sigma}{n+1-1})$.

For the "only if" part we may read the proof backwards.

2.2.3 Corollary

Let $d_i = (n+1-i)y_i$ then (d_i) are independent and $\exp(0, \sigma)$ iff (y_i) are independent and the distribution of y_i is $\exp(0, \frac{\sigma}{n+1-i})$.

2.2.4 Lemma

Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s from $\exp(0, \sigma)$ then

$$E(X_{(r)}) = \sigma \sum_{j=1}^r (n-j+1)^{-1}$$

$$\text{cov}(X_{(r)}, X_{(s)}) = \sigma^2 \sum_{j=1}^{\min(r,s)} (n-j+1)^{-2}$$

Proof

$$X_{(r)} = X_{(1)} + (X_{(2)} - X_{(1)}) + \dots + (X_{(r)} - X_{(r-1)})$$

and using the same notation as in lemma 2.2.2 we have

$$X_{(r)} = \sum_{i=1}^r Y_i$$

hence

$$E(X_{(r)}) = \sum_{i=1}^r E(Y_i)$$

$$= \sigma \sum_{i=1}^r (n+1-i)^{-1}$$

Similarly, assume $r \leq s$.

If $r=s$

$$\text{Var}(X_{(r)}) = \sum_{i=1}^r \text{Var}(Y_i)$$

$$= \sigma^2 \sum_{i=1}^r (n+1-i)^{-2}$$

since $Y_1, \dots, Y_n$ are independent.

If $r < s$

$$\text{cov}(X_{(r)}, X_{(s)}) =$$

$$= E\left(\sum_{i=1}^r Y_i - \sigma \sum_{i=1}^r (n+1-i)^{-1}\right) \left(\sum_{i=1}^s Y_i - \sigma \sum_{i=1}^s (n+1-i)^{-1}\right) =$$

$$= E\left(\sum_{i=1}^r Y_i - \sigma \sum_{i=1}^r (n+1-i)^{-1}\right) \left(\sum_{i=1}^r Y_i - \sigma \sum_{i=1}^r (n+1-i)^{-1} + \right.$$

$$\left. + \sum_{i=r+1}^s Y_i - \sigma \sum_{i=r+1}^s (n+1-i)^{-1}\right) =$$

$$= E\left(\sum_{i=1}^r Y_i - \sigma \sum_{i=1}^r (n+1-i)^{-1}\right)^2 = \sigma^2 \sum_{i=1}^r (n+1-i)^{-2}$$

We will use the following notations for the order statistics.

$$m' = (m_1, m_2, \dots, m_n) = (E(x_{(1)}), E(x_{(2)}), \dots, E(x_{(n)}))$$

$V = (V_{ij})$ an $n \times n$ matrix, where

$$V_{ij} = \text{cov}(x_{(i)}, x_{(j)})$$

2.2.5 Lemma

Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with density function $\exp(0,1)$ then

$$i) \quad m' = e'v$$

$$ii) \quad e'm = n$$

$$iii) \quad ev^{-1} = (n^2, 0, \dots, 0)$$

where $e' = (1, \dots, 1)$.

Proof

To prove (i) it is sufficient to prove $m_j = \sum_{i=1}^n V_{ij}$.

Now

$$\sum_{i=1}^n V_{ij} = \sum_{i=1}^j V_{ij} + \sum_{i=j+1}^n V_{ij} =$$

$$= \sum_{i=1}^j \sum_{k=1}^i (n-k+1)^{-2} + \sum_{i=j+1}^n \sum_{k=1}^j (n-k+1)^{-2} =$$

$$= \sum_{k=1}^j \sum_{i=k}^j (n-k+1)^{-2} + \sum_{k=1}^j \sum_{i=j+1}^n (n-k+1)^{-2} =$$

$$= \sum_{k=1}^j \sum_{i=k}^n (n-k+1)^{-2}$$

$$= \sum_{k=1}^j (n-k+1)^{-1} = m_j.$$

Proof of (ii)

$$e'm = \sum_{i=1}^n m_i = \sum_{i=1}^n \sum_{k=1}^i (n-k+i)^{-1} = \sum_{k=1}^n \sum_{i=k}^n (n-k+1) = n$$

Proof of (iii) follows immediately since

$$V^{-1} = \begin{pmatrix} n^2 + (n-1)^2, -(n-1)^2 \\ -(n-1)^2, (n-1)^2 + (n-2)^2, -(n-2)^2, \\ \vdots \\ -2^2, 2^2 + 1^2, -1^2 \\ \vdots \\ 0, -1^2, 1^2 \end{pmatrix}$$

Complete sufficient statistics for (θ, σ) .

Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with density function $\exp(\theta, \sigma)$ then the maximum likelihood statistics $\hat{\theta}, \hat{\sigma}$ for θ, σ resp. are given by

$$\hat{\theta} = x_{(1)}$$

$$\hat{\sigma} = \frac{\sum_{i=1}^n (x_i - x_{(1)})}{n}$$

The proof is trivial, and will therefore be omitted. We need the following result to prove the completeness.

2.2.6 Lemma

Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with density function $\exp(\theta, \sigma)$. Define w by

$$w = \sum_{i=1}^n (X_{(i)} - X_{(1)}),$$

then w and $X_{(1)}$ are independent and $\frac{2w}{\sigma}$ has a chi-square distribution with $(2n-2)$ degrees of freedom.

Proof

We can write w as

$$w = \sum_{i=1}^n X_{(i)} - nX_{(1)}$$

Since $X_{(r)} = X_{(1)} + (X_{(2)} - X_{(1)}) + \dots + (X_{(r)} - X_{(r-1)})$ using again the notation of lemma 2.2.2 we have

$$X_{(r)} = \sum_{i=1}^r Y_i + \theta$$

Substitution in the formula for w gives us

$$\begin{aligned} w &= \sum_{j=1}^n \sum_{i=1}^j Y_i + n\theta - nY_1 - n\theta \\ &= \sum_{j=1}^n \sum_{i=1}^j Y_i - nY_1 \end{aligned}$$

$$= \sum_{i=2}^n d_i$$

where d_i is defined in lemma 2.2.3. Since d_1 is independent of $d_2, \dots, d_n$ it follows that $x_{(1)}$ and w are independent. Since $\frac{2d_i}{\sigma}$ has a chi-square distribution with 2 degrees of freedom, for all i , $\frac{2w}{\sigma}$ has a chi-square distribution with $(2n-2)$ degrees of freedom.

2.2.7 Corollary.

$\frac{2n\hat{\sigma}}{\sigma}$ has a chi-square distribution with $(2n-2)$ degrees of freedom.

2.2.8 Theorem.

Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with density function $\exp(\theta, \sigma)$ then $T = (T_1, T_2) = (x_{(1)}, \sum_{i=1}^n x_i)$ is a complete sufficient statistic for the pair (θ, σ) .

Proof

The sufficiency follows from the fact that the joint density can be written as

$$\sigma^{-n} \exp\left\{-\frac{1}{\sigma}(T_2 - n\theta)\right\} \cdot f(T_1, \theta)$$

where

$$f(T_1, \theta) = \begin{cases} 0 & \theta > T_1 \\ 1 & \theta \leq T_1 \end{cases}$$

and therefore by the Factorization theorem T_1 and T_2 are sufficient. To show (T_1, T_2) is complete, we are required to

show:

$$\text{if } \iint f(t_1, t_2) P_{\theta, \sigma}(t_1, t_2) dt_1 dt_2 = 0$$

for all $\sigma > 0$ $\theta \geq 0$

$$\text{then } f(t_1, t_2) = 0 \quad \text{a.e.}$$

Let $w = t_2 - nt_1$ and $t = t_1$

By lemma 2.2.6 $\frac{2w}{\sigma}$ has a chi-square distribution with $2n-2$ d.f. and moreover $\frac{2w}{\sigma}$ is independent of t .

This implies

$$\int_0^\infty \int_0^\infty f(w+nt, t) \cdot w^{n-2} e^{-\frac{(w+nt)}{\sigma}} dw dt = 0$$

Hence by two-dimensional uniqueness theorem for Laplace transform

$$f(t_1, t_2) = f(w+nt, t) = 0 \quad \text{a.e.}$$

2.2.9 Lemma.

Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with density function $\exp(\theta, \sigma)$. The unique uniformly minimum variance unbiased estimate of θ based on (T_1, T_2) is given by

$$\hat{\theta} = x_{(1)} - \frac{\sum_{i=1}^n x_i - nx_{(1)}}{n(n-1)}$$

$$= \frac{nx_{(1)} - \bar{x}}{(n-1)}$$

Proof.

To prove $\hat{\theta}$ is unbiased, we have

$$E(\hat{\theta}) = \frac{nE(x_{(1)}) - E(\bar{x})}{n-1} = \frac{n(\frac{\sigma}{n} + \theta) - (\sigma + \theta)}{n-1} = \theta$$

Hence the lemma follows from theorem 2.2.8.

Writing

$$R = \frac{n(x_{(1)})^{-\theta}}{\sum_{i=1}^n x_i^{-n x_{(1)}}}$$

then the numerator and denominator are independent by lemma 2.2.6. Since

$$R = \frac{\chi^2_2}{\chi^2_{2n-2}}$$

it follows that $(n-1)R$ has a F distribution, and hence the density of R is given by

$$\begin{aligned} g(r) &= (n-1)(1+r)^{-n} & r \geq 0 \\ &= 0 & \text{otherwise.} \end{aligned}$$

2.2.10 Corollary.

The uniformly most powerful unbiased test of size α for the hypothesis that $\theta = \theta_0$, given that the sample is from an exponential distribution with σ unknown, is defined by the critical regions $r < 0$ and $r > r_\alpha$ where $\alpha = (1+r_\alpha)^{-n+1}$.

2.3.0 Transformation from exponential to uniform distribution.

Let $\underline{y} = (y_1, \dots, y_n)$ be n independent observations from some distribution F . Suppose J and L are two transformations with the following property: applying J (or L) to $\underline{y}$, say $J\underline{y} = \underline{z} = (z_1, \dots, z_{(n-1)})$ then $\underline{z}$ acts as a sample from a uniform $(0,1)$ distribution iff $\underline{y}$ is a sample from $\exp(0, \sigma)$. Hence, it is equivalent to test if the $y_i (1 \leq i \leq n)$ are a sample from $\exp(0, \sigma)$ or to test if the $z_i (1 \leq i \leq n-1)$ are uniformly distributed between $(0,1)$.

This is particularly useful when σ is unknown, since the distribution of $z_i (1 \leq i \leq n-1)$ is completely known and we can use lemma 1.3.1.

2.3.1 Definition.

Let $\{y_i | 1 \leq i \leq n\}$ be a set of non-negative real numbers.

Define

$$S_k = \sum_{i=1}^k y_i \quad k=1, \dots, n$$

Let the transformation J be defined by

$$z_j = \frac{S_j}{S_n} \quad j=1, \dots, n-1$$

and the transformation L be defined by

$$z'_j = \left(\frac{s_j}{s_{j+1}} \right)^j \quad j=1, \dots, n-1$$

we will now prove that transformations J and L have the desired property, i.e. that z'_i 's are independent and z'_i is uniformly distributed iff $\underline{y}$ is a random sample from $\exp(0, \sigma)$ under some restrictions.

Transformation J

The transformation J has been considered by Seshadri, Csörgö and Stephens and most of the proof is taken from their paper. We first need to prove some lemmas before the required result in Theorem 2.3.5.

2.3.2 Lemma (Ref. [5], [11])

Let $\{y_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with a continuous density function $p, p(y) > 0$ for all $y \in [0, \infty)$ and $E(y_i) = 1$ then writing

$$S = \sum_{i=1}^n y_i$$

$$f(y_1, \dots, y_{n-1} | S) = (n-1)! S^{-n+1}$$

iff the y_i are $\exp(0, 1)$.

Proof

The sufficiency follows from

$$f(y_1, \dots, y_{n-1}, s) = e^{-s}$$

$$f(s) = \frac{s^{n-1} e^{-s}}{(n-1)!}$$

To prove the necessity, let $p_s(s)$ denote the density of S .

The joint density of $y_1, \dots, y_{n-1}, S$ is given by

$$f(y_1, \dots, y_{n-1}, s) = \prod_{i=1}^n p(y_i)$$

Using the assumptions in the lemma

$$\frac{\prod_{i=1}^n p(y_i)}{p_s(s)} = (n-1)! s^{-n+1} \quad p_s(s) > 0$$

Let $y_2 = y_3 = \dots = y_n = 0$; we have

$$\frac{p(y_1)}{p_s(y_1)} (p(0))^{n-1} = (n-1)! y_1^{-n+1}$$

Thus

$$\frac{p(s)}{p_s(s)} \cdot (p(0))^{n-1} = (n-1)! s^{-n+1}$$

But this with the above assumption implies

$$\frac{p(y_1)}{p(0)} \times \frac{p(y_2)}{p(0)} \times \dots \times \frac{p(y_n)}{p(0)} = \frac{p(y_1 + \dots + y_n)}{p(0)}$$

Let $\phi(x) = \frac{p(x)}{p(0)}$.

We have
$$\prod_{i=1}^n \phi(y_i) = \phi(y_1 + \dots + y_n)$$

whose solution is $\phi(x) = e^{ax}$ where a is an arbitrary constant.

Hence

$$p(y_i) = p(0)e^{ay_i} \quad i=1, \dots, n$$

Since $p(y)$ is a density function with mean equal to 1,

$$p(y_i) = e^{-y_i}$$

This completes the proof of the lemma.

2.3.3 Corollary

Let $D_i = y_i/s$ ($1 \leq i \leq n$). Then the D_i are jointly distributed as n spacings generated by $(n-1)$ order statistics of a random sample from uniform $(0,1)$ iff the y_i are $\exp(0,1)$.

Proof

From lemma 2.3.2 we have

$$f(D_1, \dots, D_{n-1} | S) = (n-1)!$$

iff the y_i are $\exp(0,1)$ distributed. The conditional density being independent of s implies

$$f(D_1, \dots, D_{n-1}) = (n-1)!$$

This is the density of the spacings of $n-1$ observations from a uniform $(0,1)$ distribution.

2.3.4 Corollary.

Let

$$z_r = \sum_{i=1}^r D_i = \frac{\sum_{i=1}^r Y_i}{s} \quad 1 \leq r \leq n-1$$

then the z_r acts like $(n-1)$ ordered statistic from uniform $(0,1)$ iff Y_i are $\exp(0,1)$ distributed.

2.3.5 Theorem.

Let $\{Y_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with a continuous density function $p, p(y) > 0$ for all $y \in [0, \infty)$ and $E(Y_i) = \sigma > 0$, write

$$\begin{aligned} \text{i)} \quad s &= \sum_{i=1}^n Y_i \\ \text{ii)} \quad z_r &= \frac{\sum_{i=1}^r Y_i}{s} \quad 1 \leq r \leq n-1 \end{aligned}$$

Then the z_i act like $(n-1)$ ordered r.v. from uniform $(0,1)$ iff Y_i are $\exp(0, \sigma)$ distributed.

The proof follows immediately by letting $x_i = \frac{Y_i}{\sigma}$, and applying Corollary 2.3.4 to x_i .

Transformation L

2.3.6 Theorem.

Let $\{Y_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with a continuous

density function $p, p(y) > 0$ for all $y \in [0, \infty)$, and $E(Y_i) = \sigma > 0$.

$$\text{Let} \quad S_k = \sum_{i=1}^k Y_i \quad k = 1, \dots, n$$

$$\text{and} \quad Z'_j = \frac{S_j}{S_{j+1}} \quad j=1, \dots, n-1$$

then Z'_j are independent and uniformly $(0,1)$ distributed iff Y_i are $\exp(0, \sigma)$ distributed.

Proof of Sufficiency

The joint density of Y_i ($i=1, \dots, n$) is given by

$$\begin{aligned} g(y_1, \dots, y_n) &= \frac{1}{\sigma^n} \exp\left\{-\frac{1}{\sigma} \sum_{i=1}^n y_i\right\} & y_i > 0 \quad \forall i \\ &= 0 & \text{otherwise} \end{aligned}$$

$$\text{Let} \quad Y_1 = Y_1$$

$$Z'_m = \frac{\sum_{i=1}^m y_i}{\sum_{i=1}^{m+1} y_i} \quad m=1, \dots, n-1$$

$$\text{or equivalently} \quad Y_1 = Y_1$$

$$\begin{aligned} Y_2 &= \frac{Y_1}{Z'_1} - Y_1 \\ &\vdots \end{aligned}$$

$$y_n = \frac{y_1}{w_1 w_2^{1/2} \dots w_{n-1}^{1/n-1}} - \frac{y_1}{w_1 w_2^{1/2} \dots w_{n-2}^{1/n-2}}$$

Hence the jacobian of the transformation, is given by

$$|J| = \begin{vmatrix} 1 & , & \\ a_{21} & , & -\frac{y_1}{w_1^2} & , \\ \vdots & & \vdots & \\ a_{n1}, a_{n2} & , \dots & , & -\frac{y_1}{(n-1) w_1 w_2^{1/2} \dots w_{n-1}^{n/n-1}} \end{vmatrix}$$

$$= \frac{y_1^{n-1}}{(n-1)! w_1^n \dots w_{n-1}^{n/n-1}} = \frac{y_1^{n-1}}{(n-1)! c^n} \quad (\text{say})$$

Now
$$\sum_{i=1}^n y_i = \frac{y_1}{w_1 w_2^{1/2} \dots w_{n-1}^{1/n-1}} = \frac{y_1}{c}$$

Hence the joint density of $w_1, \dots, w_{n-1}$ is given by

$$g(w_1, \dots, w_{n-1}) = \frac{1}{(n-1)! c^n \sigma^n} \int_0^\infty e^{-\frac{y_1}{\sigma c}} y_1^{n-1} dy_1$$

$$= \frac{\sigma^{n-1} (c')^{n-1}}{(n-1)! c_\sigma^n} \int_0^\infty e^{-\frac{y_1}{\sigma c'}} \left(\frac{y_1}{\sigma c'}\right)^{n-1} dy_1$$

$$g(w_1, \dots, w_{n-1}) = 1 \quad .$$

To prove the necessity

$$W_{n-1}^{1/n-1} = \frac{\sum_{i=1}^{n-1} y_i}{\sum_{i=1}^n y_i} = \frac{y_1}{s_n} + \dots + \frac{y_{n-1}}{s_n}$$

$$= \sum_{i=1}^{n-1} D_i = Z_{n-1} \quad \text{where } D_i = \frac{y_i}{s_n}$$

hence the density of Z_n is given by

$$g(Z_n) = n Z_n^{n-1} \quad 0 \leq Z_n \leq 1$$

$$= 0 \quad \text{otherwise}$$

Hence Z_{n-1} is distributed as n th ordered statistic, from a sample of size n from a uniform $(0,1)$ distribution

$$W_{n-1}^{1/n-1} \cdot W_{n-2}^{1/n-2} = \sum_{i=1}^{n-2} D_i = Z_{n-2}$$

and is easy to check that Z_{n-2} have the same distribution as the $(n-1)^{th}$ ordered statistic from a sample of size n from uniform $(0,1)$ distribution, and so on. Hence the result follows from theorem 2.3.5.

One disadvantage with transformation $J(L)$ is that $Z_j(Z'_j)$ depends on the order of y_i , i.e. the test is not invariant under permutation of the sample. However, this problem can be solved in the following way. Recalling theorem 2.2.2 (p. 22) and corollary 2.2.3 (p. 23) we have: Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s with distribution function $\exp(0, \sigma)$. Let $Y_i = (n+1-i)(X_{(i)} - X_{(i-1)})$ $i=1, \dots, n$ with $X_{(0)} \equiv 0$, then Y_i 's are independent and $\exp(0, \sigma)$ distributed. This transformation we will call G . Hence

$$\underline{Z} = JG\underline{y} = J'\underline{y}$$

$$\underline{Z} = LG\underline{y} = L'\underline{y}$$

and the transformation J' and L' are invariant under permutation of the sample.

2.4.0 The W-Statistic.

Generalized Least Squares.

We write a linear model in the form

$$\underline{y} = \underline{X}\underline{\theta} + \underline{\varepsilon}$$

where $\underline{y}$ is an $(n \times 1)$ vector of observations

$\underline{X}$ is an $(n \times k)$ matrix of known coefficients

$\underline{\theta}$ is an $(k \times 1)$ vector of parameters

$\underline{\varepsilon}$ is an $(n \times 1)$ vector of "errors",

a r.v. with

$$E(\underline{\varepsilon}) = \underline{0}$$

$$\text{Var}(\underline{\varepsilon}) = \sigma^2 \underline{I}_n$$

The least squares method requires that we minimize the scalar sum of squares

$$S = (\underline{y} - \underline{X}\underline{\theta})' (\underline{y} - \underline{X}\underline{\theta})$$

for variation in the components of θ .

The generalized case is the same except that

$$\text{Var}(\underline{\varepsilon}) = \sigma^2 \underline{V}$$

where $\underline{V}$ is an $(n \times n)$ completely specified matrix.

Lloyd [9] considers the special case of estimating the location and scale parameters from the order statistics.

$$\text{Let } x_{(r)} = \frac{y_{(r)} - \mu}{\sigma} \quad r=1, \dots, n$$

$$\text{and } E(\underline{X}) = \underline{\alpha}$$

$$\text{Var}(\underline{X}) = \underline{V}$$

$$\underline{y} = \mu \underline{e}' + \underline{\alpha}$$

Then the generalized least squares estimate is given by

$$\hat{\underline{\mu}} = \frac{\underline{-\alpha}' \underline{V}^{-1} (\underline{e\alpha}' - \underline{\alpha e}') \underline{V}^{-1} \underline{y}}{\Delta}$$

$$\hat{\sigma} = \frac{\underline{e}' \underline{V}^{-1} (\underline{e\alpha}' - \underline{\alpha e}') \underline{V}^{-1} \underline{y}}{\Delta}$$

where

$$\Delta = \{ (\underline{e}' \underline{V}^{-1} \underline{e}) (\underline{\alpha}' \underline{V}^{-1} \underline{\alpha}) - (\underline{e}' \underline{V}^{-1} \underline{\alpha})^2 \}$$

Shapiro and Wilks [12] adapted this result for the exponential case.

Let $y = (y_{(1)}, \dots, y_{(n)})$ denote an ordered sample from $\exp(\theta, \sigma)$, and let

$$y_{(i)} = \theta + \sigma(X_{(i)}) \quad i=1, \dots, n$$

Using the notation of lemma 2.2.5 (p. 25),

$$\begin{aligned} \hat{\sigma} &= \frac{\underline{eV}^{-1} (\underline{em}' - \underline{me}') \underline{V}^{-1} \underline{y}}{\underline{e}' \underline{V}^{-1} \underline{m}' \underline{V}^{-1} \underline{m} - (\underline{e}' \underline{V}^{-1} \underline{m})^2} \\ &= \frac{n}{n-1} (\bar{y} - y_{(1)}) \end{aligned}$$

2.4.1 Definition.

The W-statistic is defined by

$$W = \frac{n}{n-1} \frac{(\bar{Y} - Y_{(1)})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

where W is to be used as a two-tailed statistic.

Note that W is a ratio of two estimates of variance, except for a factor of n. W can also be written as

$$w = \frac{((1-n)Y_{(1)} + \sum_{i=2}^n Y_i)^2}{n(n-1) \sum_{i=1}^n (Y_i - \bar{Y})^2} = \frac{(\sum_{i=1}^n a_i Y_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

where $\sum_{i=1}^n a_i = 0$ and $\sum_{i=1}^n a_i^2 = 1$

Properties of W-statistic

2.4.2 Lemma.

W is scale and location invariant. The proof follows immediately from the definition.

2.4.3 Lemma.

The maximum value of W is 1. The minimum value of w is $(n-1)^{-2}$.

Proof.

Since W is location invariant we assume $\bar{y} = 0$.

That W is less than 1, follows from

$$(\sum a_i y_i)^2 \leq \sum a_i^2 \sum y_i^2, \text{ and since}$$

$$\sum_{i=1}^n a_i^2 = 1$$

the equality holds when $y_i = a_i$ $i=1, \dots, n$

To prove minimum, it is sufficient to maximize

$$\sum_{i=1}^n y_i^2 \text{ subject to the constraint } \sum_{i=1}^n y_i = 0 \text{ and } \sum_{i=1}^n a_i y_i = 1,$$

since W is scale invariant. Since all vectors satisfying

$$\sum_{i=1}^n y_i = 0, \quad \sum_{i=1}^n a_i y_i = 1 \text{ is a convex region and } \sum_{i=1}^n y_i^2 \text{ is}$$

a convex function the maximum must occur at one of the $(n-1)$ vertices of the region. These are

$$\frac{n-1}{na_1}, \frac{-1}{na_1}, \dots$$

$$\begin{array}{ccccccc} \frac{(n-2)}{n(a_1+a_2)} & , & \frac{(n-2)}{n(a_1+a_2)} & , & \frac{-2}{n(a_1+a_2)} & , & \dots, \frac{-2}{n(a_1+a_2)} \\ \vdots & & & & & & \vdots \\ & & \frac{1}{n(a_1 + \dots + a_{n-1})} & , & \dots & & \frac{n-1}{n(a_1 + \dots + a_{n-1})} \end{array}$$

It is not difficult to check that the maximum occurs in the last row.

$$(a_1 + \dots + a_{n-1}) = \frac{-1}{\sqrt{n(n-1)}}$$

$$\sum_{i=1}^n y_i^2 = (n-1) \left(\frac{n(n-1)}{n^2} \right) + \frac{(n-1)^2 n(n-1)}{n^2} = (n-1)^2$$

Hence the minimum is equal to $(n-1)^2$, and the lemma is proven.

In the next lemma we will use the following result by Basu [2].

Let the family of possible distributions of X be $\mathcal{P} = \{P_v, v \in \Omega\}$. Let T be a sufficient statistic for v , and suppose that the family $\mathcal{P}^T$ of distributions of T is boundedly complete. Then a statistic V is independent of T for all $v \in \Omega$ iff the distribution of V does not depend on v .

2.4.4 Lemma.

Let $\{X_i | 1 \leq i \leq n\}$ be i.i.d. r.v.'s from $\exp(\theta, \sigma)$ then the w -statistic is independent of $X_{(1)}$ and $\bar{X}$.

Proof

Follows immediately from Basu result.

2.4.5 Corollary.

R and W are independent where R is defined in 2.2.10. Hence we can combine R and W as a test for $(H_1: \theta_0, \sigma)$. One way for this compounding, is using the following result.

If X is uniformly $(0, 1)$ distributed then $-2\ln X$ has a chi-square distribution with 2 d.f. Hence

$$T_1 = -2\ln(\psi_R \cdot \psi_W)$$

where ψ_R and ψ_W are significance levels actually obtained by R and W statistics. Then T_1 has a chi-square distribution with 4 d.f.

CHAPTER III

3.1.0 Introduction.

One disadvantage of using the EDF-statistic has been the need for extensive tables. Stephens [13] has suggested that for each test statistic, T say, a simple modification, T^* , is given and T^* is compared with the given percentage points. We will discuss this modification in section 3.2.0. In section 3.3.0 we will consider another approximation to the percentage points, such that standard chi-square tables can be used.

In section 3.4.0 we will discuss the power of the EDF-statistics and compare these with Shapiro and Wilks' W -statistic.

3.2.0 Modification of EDF-statistic.

In the literature, extensive tables exist for most of the EDF-statistics. Stephens has suggested a method whereby these tables may be replaced by shorter ones.

However, the modification T^* is only an approximation of T_n . If the approximate point is calculated for the significance level α , and the true level is α' , the error $|\alpha - \alpha'|$ is negligibly small if $n \geq 8$ for all practical purposes. The reason the tables for T^* can be made shorter is that they do not depend on n .

If the hypothesised distribution is known, for example,

after using the transformation J or L, the following modification was suggested:

Statistic T	Modified T*
D_n^+	$D^{+*} = D_n^+(\sqrt{n}+0.12+0.11/\sqrt{n})$
D_n^-	$D^{-*} = D_n^-(\sqrt{n}+0.12+0.11/\sqrt{n})$
D_n	$D^* = D_n(\sqrt{n}+0.12+0.11/\sqrt{n})$
V_n	$V^* = V_n(\sqrt{n}+0.155+0.24/\sqrt{n})$
W_n^2	$W^* = (W_n^2 - 0.4/\sqrt{n} + 0.6/n^2)(1.0 + 1.0/n)$
U_n^2	$U^* = (U_n^2 - 0.1/\sqrt{n} + 0.1/n^2)(1.0 + 0.8/n)$
A	$A^* = A \quad \text{for all } n \geq 5$

The percentage points for T* are given in Table I.

The percentage points for the EDF-statistics, when the samples were taken from an exponential distribution with unknown mean, have been found, using the Monte Carlo method. Lillefors [8] did this first for the Kolmogorov-Smirnov statistic, and Stephens [13] completed the tables for the other EDF-statistics.

For those values of n where the exact values of

EDF-statistics are known, the simulated values agree very closely with both Lillefors' and Stephens' results.

Similarly, for the case when $F(x)$ is completely specified, Stephens [13] suggested the approximation $\tilde{T}^*$, to simplify the tables. In this case the modifications are given as follows:

Statistic T	Modification $\tilde{T}^*$
$\tilde{D}_n$	$\tilde{D}^* = (\tilde{D}_n - 0.2/n) (\sqrt{n+0.26} + 0.5/\sqrt{n})$
$\tilde{V}_n$	$\tilde{V}^* = (\tilde{V}_n - 0.2/n) (\sqrt{n+0.24} + 0.35/\sqrt{n})$
$\tilde{W}_n^2$	$\tilde{W}^* = \tilde{W}_n^2 (1 + 0.16/n)$
$\tilde{U}_n^2$	$\tilde{U}^* = \tilde{U}_n^2 (1 + 0.16/n)$
$\tilde{A}_n$	$\tilde{A}^* = \tilde{A}_n (1 + 0.6/n)$

The percentage points for $\tilde{T}^*$ are given in Table II.

2.3.0 Chi-square approximation.

A sufficiently good approximation of $\tilde{T}^*$ can be found in the form

$$\tilde{T}^*(\alpha) = a + b\chi_{\nu}^2(\alpha)$$

where the constants a , b and ν are obtained by matching the moments of $\tilde{T}^*$. The moments of $\tilde{T}^*$ were obtained by the

Monte-Carlo method. Using this method, a standard chi-square table can be used for all EDF-statistics. In Table III we give the first four moments of $\tilde{T}^*$. The constants a , b and v are given by:

$$\tilde{D}^*(\alpha) = 0.017 + 0.0343\chi_{20}^2(\alpha)$$

$$\tilde{V}^*(\alpha) = -0.336 + 0.0295\chi_{50}^2(\alpha)$$

$$\tilde{W}^*(\alpha) = 0.0460 + 0.0466\chi_1^2(\alpha)$$

$$\tilde{U}^*(\alpha) = 0.0265 + 0.0266\chi_2^2(\alpha)$$

where $\chi_v^2(\alpha)$ is the upper tail percentage point at level α , of a chi-square distribution with v d.f.

Although this appears to be an approximation of an approximation, the results are very accurate. For a comparison, see Stephens [13].

3.3.0 Power of the EDF-statistic.

In Chapters I and II we have discussed five different ways of using EDF-statistics for testing for exponentiality, i.e.

- i) Integral transformation
- ii) J transformation
- iii) L transformation
- iv) J' transformation
- v) L' transformation

We have also considered the

- i) W-statistic
- ii) T_1 -statistic

We now have the following problems:

- 1) How does the power compare between the different EDF-statistics?
- 2) What transformation gives the greatest power?

To find the power, we used the Monte-Carlo method and selected 5,000 random samples from the following distributions:

- a) Half normal (i.e., absolute value of samples from $N(0,1)$)
- b) Half Cauchy
- c) Uniform (0,1)
- d) $F(x) = 1-(1-x)^{3/2}$
- e) $F(x) = 1-(1-x)^3$
- f) Chi-square (1-df)

and then tested for exponentiality. Table IV gives the results of this study. The study indicates that the Shapiro-Wilks W-statistic has greater power than any of the EDF-statistics, with the exception of the case where the sample comes from a chi-square distribution. As expected $\tilde{W}_n^2$ has greater power than V_n , since $\tilde{W}_n^2$ is a "sum" of discrepancies. More complete tables will be published by Stephens.

TABLE 1[†]

Modified EDF-Statistic, When $F(x)$ Completely Known

T*	Percentage Points for T*				
	15	10	5	2.5	1
D^{+*}, D^{-*}	.973	1.073	1.224	1.358	1.518
D^*	1.138	1.224	1.358	1.480	1.628
V^*	1.537	1.620	1.747	1.862	2.001
W^*	0.284	0.347	0.461	0.581	0.743
U^*	0.131	0.152	0.187	0.221	0.267
A^*	1.61	1.933	2.492	3.070	3.857

[†]Source Stephens [14]

TABLE 2[†]

Modified EDF-Statistic for Test of Exponentiality
Scale Parameter Unknown

T*	Percentage Points for T*				
	15	10	5	2.5	1
D*	0.926	0.990	1.094	1.190	1.308
V*	1.445	1.527	1.655	1.774	1.910
W*	0.149	0.177	0.224	0.273	0.337
U*	0.112	0.130	0.161	0.191	0.230
A*	0.922	1.078	1.341	1.606	1.957

[†]Source Stephens [14]

TABLE 3[†]

Estimated Moments of Null Distributions:
Monte Carlo Results, Not Smoothed

		Statistic Moments			
	n	m'_1	m'_2	m'_3	m'_4
$\sqrt{n}D$	6	0.673	0.485	0.375	0.308
	10	.687	.507	.400	.337
	20	.703	.531	.431	.375
	50	.712	.545	.448	.395
	100	.720	.557	.462	.412
w^2	6	0.094	0.013	0.0024	0.0006
	10	.094	.013	.0026	.0007
	20	.095	.014	.0029	.0009
	50	.094	.013	.0027	.0008
	100	.094	.013	.0027	.0007
$\sqrt{n}V$	6	1.084	1.232	1.466	1.824
	10	1.106	1.285	1.563	1.989
	20	1.136	1.355	1.696	2.225
	50	1.154	1.400	1.780	2.373
	100	1.169	1.433	1.843	2.483
U^2	6	0.074	0.007	0.001	0.0002
	10	.074	.007	.001	.0002
	20	.074	.007	.001	.0002
	50	.073	.007	.001	.0002
	100	.073	.007	.001	.0002

[†] Source Stephens [13]

TABLE 4A

Power Comparisons; Test for Exponentiality, θ Unknown. The Table Gives the Percentage of Samples Significant, When the Population is as Shown, and the Sample Size is 20. The Test is at the 10% Level. The EDF-Statistics are Calculated Using Studentized Integral Transformation.

	$\tilde{D}_{20}^2$	$\tilde{W}_{20}^2$	$\tilde{V}_{20}^2$	$\tilde{U}_{20}^2$	$\tilde{A}_{20}$	T_1	W
Half Normal	27.4	30.5	24.5	25.8	27.0	8.9	30.2
Half Cauchy	66.0	68.6	58.4	60.0	69.5	51.0	72.8
Uniform	67.0	78.7	77.0	73.0	77.0	14.0	84.3
$F(X) = 1-(1-X)^{3/2}$	42.9	52.0	46.2	44.0	49.0	10.0	60.8
$F(X) = 1-(1-X)^3$	20.0	22.7	19.0	19.5	20.0	6.0	24.5
Chi-Square 1.df.	57.6	62.0	50.8	53.4	73.6	26.0	40.6

TABLE 4B

Power Comparison; Test for Exponentiality, θ Unknown. The Table Gives the Percentage of Samples Significant When the Population is as Shown, and the Sample Size 20. The Test is at 10% Level. The EDF-Statistics are Calculated Using Transformation L.

	D_{20}^2	W_{20}^2	V_{20}^2	U_{20}^2	A_{20}	W
Half Normal	13.0	12.0	20.0	21.0	10.2	30.2
Half Cauchy	42.8	44.0	46.0	47.0	54.0	72.8
Uniform	26.0	28.4	59.5	57.6	25.6	84.3
$F(X) = 1-(1-X)^{3/2}$	16.7	16.8	34.4	34.0	14.3	60.8
$F(X) = 1-(1-X)^3$	10.5	10.4	15.8	16.4	9.0	24.5
Chi-Square 1.df.	37.9	38.6	47.0	48.8	64.5	40.6

TABLE 4C

Power Comparison. Test for Exponentiality, θ Unknown. The Table Gives the Percentage of Samples Significant, When the Population is as Shown, and the Sample Size 20. The Test is at 10% Level. The EDF-Statistics are Calculated Using Transformation L' .

	D_{20}^2	W_{20}^2	V_{20}^2	U_{20}^2	A_{20}	W
Half Normal	23.0	26.0	16.0	16.0	25.0	30.2
Half Cauchy	26.0	29.0	18.0	20.0	48.0	72.8
Uniform	60.0	65.0	38.0	38.0	66.0	84.3
$F(X) = 1-(1-X)^{3/2}$	38.0	41.0	24.0	25.0	42.0	60.8
$F(X) = 1-(1-X)^3$	17.0	18.0	13.0	14.0	18.0	24.5
Chi-Square 1.df.	63.0	67.0	43.0	43.0	76.0	40.0

10. Marshall, A.W. "The small-sample distribution of nW_n^2 ",
Ann. Math. Stat. 29 (1958) 307-309.
11. Seshadri, V., Gsörgö, M. and Stephens, M.A. "Test for the
exponential distribution using Kolmogorov-type statis-
tics", J.R. Stat. Soc. B. (1969) 499-509.
12. Shapiro, S.S. and Wilks, M.B. "An analysis of variance
tests for the exponential distribution", Bell Telephone
Laboratories (1969) Case 20878.
13. Stephens, M.A. "Kolmogorov-type tests for exponentiality
when the scale parameter is unknown", Technical Report
No. 154 (1970) Department of Statistics, Stanford
University.
14. Stephens, M.A. "EDF-statistics for goodness-of-fit", un-
published paper.
15. Smirnov, N.V. "On the estimation of the discrepancy between
empirical curves of distribution for two independent
samples", Bull. Math. Univ. Moscow 2 (No. 2) (1939)
3-14.
16. Smirnov, N.V. "On the distribution of the W^2 criterion of
Von Mises", Rec. Math. (N.S.) 2 (1937) 973-993.
17. Von Mises, R. "Wahrscheinlichkeitsrechnung", Leipzig-Wien
(1931).
18. Watson, B.G. "Goodness-of-fit tests on a circle", Biometrika
48 (1961) 109-114.

REFERENCES

1. Anderson, T.W. and Darling, O.A. "Asymptotic theory of certain 'goodness-of-fit' criteria based on stochastic processes", Ann. Math. Stat. 23 (1952) 193-212.
2. Basu, D. "On statistics independent of a complete sufficient statistic", Sankhya 15 (1955) 377-380.
3. Cramér, H. "On the composition of elementary errors", Skandinavisk Aktuarie Tidsskrift, 11 (1928) 13-74, 141-180.
4. Epstein, B. and Sobel, M. "Life testing I", J. Amer. Stat. Assn. 48 (1953) 486-502.
5. Ghurye, S.G. "Characterization of some location and scale parameter families of distribution", Cont. Prob. Stat. (1960) 203-214.
6. Kolmogorov, A.N. "Sulla determinazione empirica di una legge di distribuzione", Giorn. dell'Institut. Degli All. 4 (1933) 83-91.
7. Kuiper, N.H. "Test concerning random points on a circle", Proc. Koninkl. Nederl. Akad. Van. Wetenschappen, Series A (1970) 383-397.
8. Lilliefors, H.W. "On the Kolmogorov-Smirnov tests for the exponential distribution with mean unknown", J. Amer. Stat. Assn. 64 (1969) 387-389.
9. Lloyd, E.H. "Least squares estimation of location and scale parameters using order statistics", Biometrika 51 (1952) 484-487.

B30047